Understanding Mobile Information Supply: Studying the Amount of Textual Information Smartphones Provide

David Petersen and Matthias Böhmer
Technische Hochschule Köln, Cologne, Germany
{david.petersen,matthias.boehmer}@th-koeln.de

ABSTRACT
Smartphones became powerful tools for interacting with information anytime and anyplace. This paper investigates mobile information supply by assessing smartphones as information channels. We apply concepts of information theory for estimating the amount of information that smartphones supply. We contribute a measure of mobile information supply and describe a large-scale study analyzing data from more than 790 participants. Our main findings are that our participants' smartphones supply about 13 bytes of information per second and 1143 bytes per session. Further, within device sessions mobile information supply peaks at the beginning and the end. Our data also shows differences between sessions regarding the ratio of supplying old content vs. new content.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation: Miscellaneous

Author Keywords
Mobile apps, information supply, measure, large scale study.

INTRODUCTION
Smartphones are ubiquitous and people became used to making use of them as multifunctional tools. For many tasks they support us as information appliances for retrieving information anytime and anyplace. More and more mobile apps are confronting us with a growing number of notifications and a vast amount of information from various sources, like news, social networks and communication.

Since people revisit their apps [11] redundant information is prevalent. Examples can be found easily: people check mails repeatedly without seeing any new messages or people see the same news post on different social networks like Twitter and Facebook. Also, people repeatedly check different sources like social networks or news streams but no new content is available — known as checking habit [16]. Understanding how people interact with the information at hand can be used for improving mobile information supply, e.g. when designing mobile apps. We set forth to answer important questions with regard to mobile information supply, like how much information do mobile devices supply throughout the day? Do users repeatedly see redundant information when checking for new information? How does information supply correlate with general device usage?

Our work is motivated by Pirolli and Card's [20] early work on information foraging theory. The theory describes how people search for information. For instance, it was found that while searching for a certain piece of information on the web people browse on a specific website until the information gained by the website diminishes [10]. While the theory was applied to users searching for information on the web on stationary computers, we study if the theory also describes information foraging on smartphones — which are different to stationary computers in many concerns (e.g. size, mobility, contextual usage, attention span).

This paper has a twofold contribution: First, we describe a measure for quantifying the amount of information supplied by smartphones and present an implementation. Second, we present results of a large-scaled study using this measure to understand how mobile information is supplied.

RELATED WORK

Push and Pull of Mobile Information
On smartphones we can distinguish between information being pulled from their sources (e.g. user opening a social app for retrieving peers' status updates) vs. being pushed by their sources (e.g. notifications about content from news channels).

Pulled information mainly relates to users launching apps. Böhmer et al. [2] present a large-scale study on mobile app usage reporting that news apps are more frequently used in the morning, while game applications are more used at night, while communication apps are equally important throughout the day. Also Karikoski et al. [12] report on usage of mobile applications to be dependent on user's place-related context. Brown et al. [3] study smartphone usage in context leveraging video recording. They find that interaction with smartphones is weaved into other activities. In particular they find information search to be a time and attention-consuming activity requiring focal attention most of the time. Van Berkel et al. [28] study gaps in smartphone interaction and present an approach for identifying coherent smartphone usage sessions. They propose to use a 45 seconds threshold for gaps before considering device usage sessions to be separate. Regarding information retrieval on smartphones, i.e. content pulled from the web,

NordiCHI'18, September 29-October 3, 2018, Oslo, Norway

ISBN 978-1-4503-6437-9/18/09…$15.00

DOI: https://doi.org/10.1145/3240167.3240205

Carrascal and Church [4] find that users interact with more apps for longer durations if they use mobile search engines.

Push notifications are another large source of information on mobile devices. Pielot et al. [18] report that on average users receive about 64 notifications daily (mostly messages and emails). Shirazi et al. [21] find that in general notifications are disruptive, messenger notifications are considered most important by users, but even important notifications do not imply immediate attention. Mehrotra et al. [15] analyze how users perceive smartphone notifications: users rate the perceived disruption based on the relationship between the person who triggered the notification (if any), the user's current task, as well as that task's progress and complexity.

While there is a growing understanding on mobile application usage, information pull and push notifications, so far there is no quantitative measure for mobile information supply.

Habits and Patterns around Information Needs

Oulasvirta et al. [17] as well as Ferreira et al. [9] report on app interactions to happen in short bursts. While Oulasvirta et al. [17] discuss the fragmented nature of attention to mobile devices within bursts of 4-8 seconds, Ferreira et al. [9] find that such micro interactions happen frequently, mainly when users are alone. Böhmer et al. [2] found that when unlocking the smartphone, people often use only one single application until the device will be set back into standby state. Oulasvirta et al. [16] found that people's behavior of briefly and repetitively checking their phones for new pieces of information is habit forming. This serves as an entry point for using the smartphone further than the sole inspection of new dynamic content. The authors describe that due to the habit people also regularly check their phones even if no new content is available. Also Banovic et al. [1] provide evidence for such very short mobile interactions and are able to distinguish glance, review and engage sessions. Taking this further, Jones et al. [11] investigate users' patterns for revisiting smartphone apps. They argue that the habitual patterns of mobile phone use are not driven by technology characteristics, but rather by the users' needs for specific services and information.

However, so far it is unknown how much information smartphones supply during app usage. Further, the relation between new content and such already seen before has not been studied.

Measuring Information

Lord et al. [13] study smartphones as information and entertainment devices. Looking into the connectivity patterns they find an intensive data demand of information-driven mobile applications for supporting users in their daily practices. In the same vein, Widdicks et al. [29] found that watching, online dating, listening, social networking and communication are the top most data demanding app categories. However, it remains unclear how user interaction relates to data transfer, e.g. data transferred in the background does not have to be presented to a user at all. In contrast to that, Donderi [8] and Tuch et al. [27] investigate the amount of information conveyed by a piece of content shown to a user. Donderi [8] measures the visual complexity of an image by compressing it. He finds out that the information content can be measured by evaluating the logarithm of the zip-compressed file size of an image. Based on this finding, Tuch et al. [27] are using the jpeg algorithm to determine the visual complexity of website screenshots. They show that the file size of a jpeg image correlates with test subjects rating for visual clutter, well organized (negatively correlated) and overloaded. We extend this approach by contributing a text based measure, because on smartphones text widgets are most used view in mobile graphical interfaces (35%), as Sahami et al. presented a study on mobile application UI layouts [22].

While mobile information need as well as habits and patters have been studied with regard to different aspects as described above, so far there has been no study of smartphone usage looking into the amount of information supplied by mobile devices. This paper addresses this gap.

MEASURING MOBILE INFORMATION SUPPLY

Apart from modalities like vibration or ringtones the display represents the primary user interface on mobile devices. It does not only show content, but also allows for direct interaction with presented elements through touch input. We consider smartphones as information channels conveying information from various sources to their users through their displays.

Text is the most prominent element of mobile graphical user interfaces [22]. Thus, we focus our analysis on information presented as textual content. However, this approach is not limited to text widgets containing textual information, but also incorporates other widgets like buttons and input fields. Our approach does not take into account information conveyed through pictures, which constitute 16% of screen widgets [22], or videos. However, many images have additional meta texts for making them accessible to blind people. This information is covered in our approach.

Concept for Measuring Mobile Information Supply

We follow an information theoretic approach for determining the amount of information provided to a user. The information amount of text-based content can easily be quantified by applying information theory and eliminating redundancy, e.g. by searching for duplicate strings. Shannon specifies the term entropy as a measure for the information amount of transmitted messages [24]. Entropy is primarily based on the probabilities of single characters or rather the uncertainty of a message content. When working with redundant characters, these could be represented with a shorter code to identify them. Compression algorithms apply this theory for reducing amount of information required for encoding text.

We apply this information theoretic approach to text conveyed through smartphone displays and use compression for extracting the amount of information (like [8, 27]).

In the following we denote a screen as visual output generated by a smartphone display at a specific point in time. Whenever the shown content changes (e.g. a notification shows up) we consider this as a new screen. Assuming that two subsequent screens n and $n+1$ show texts $text_n$ and $text_{n+1}$. The amount of new information on the second screen contributes can be estimated as follows: The first text $text_n$ is compressed (here: function *comp*). The remaining byte size of the compressed result is $comp(text_n)$. Then the first and second texts will be concatenated (here: $text_n + text_{n+1}$) and compressed: $comp(text_n + text_{n+1})$. The difference of both results $comp(text_n + text_{n+1}) - comp(text_n)$ is a measure for the amount new information contributed by $text_{n+1}$ which has not been available before: Hence, this is the new information supplied to the user by the second screen. This amount of new information Δi_n a new screen n provides with regard to all the information which has been available on previous screens $1,\ldots,n-1$ before can be estimated:

$$\Delta i_n = comp(\sum_{i=1}^{n} text_i) - comp(\sum_{i=1}^{n-1} text_i)$$

Note, that this measure has an abstract nature with byte unit. While the value has its origin in text as natural language it cannot be transferred back into written text, due to compression. This measure, which we call information supply, is an abstract metric for the amount of new information which is presented to users by their smartphones.

Implementation of Information Measure

We implemented this measure for the Android mobile operating system leveraging accessibility services[1]. As they have been designed to assist users with disabilities with applications like screen readers, they provide full access to the text-based content being presented on screen. This includes the content of the current app, as well as the content provided by the operating system itself, like notification popups. In addition, they also provide events for whenever the content presented on the current screen changes. This is the case when a ui element gets visible to the user or when an already visible element changes content. Hence, we were able to use this service for capturing series of screens $1,\ldots,n$ and measuring the amount of information supplied by the current screen Δi_n. We however cannot identify if the displayed information is pulled or pushed nor can we take into account whether the user is actually looking at the data presented by the smartphone or not.

For extracting the amount of information a text-based compression algorithm must not only consider the single characters but also be able to identify the exact same character sequences. Therefore, for our implementation we used the Deflate algorithm as it is using a Huffman coding for the character probabilities and a LZ77 algorithm to identify redundant character strings [7]. The latter is carried out by searching past input values for redundant character strings within a sliding window. When a match is found, not the string is saved redundantly but a reference to the old one. In our implementation this sliding window for scanning redundant strings has a size of 32 KB, which complies to 32.768 characters. This means, that redundant character strings can be only found within the last 32 KB raw data. After implementing a proof of concept, we had to limit the total history of text we were able to store (i.e. $\sum_{i=1}^{n} text_i$) to about 16 MB in total, due to storage and computational constraints of mobile apps. Thus only the last 16 MB of raw text data can be considered when determining the information supply. That means an already seen $text_x$ is not identifiable as redundant and would be considered as new information, as soon as the text data that shows up after $text_x$ exceeds 16 million characters.

[1] https://developer.android.com/guide/topics/ui/accessibility/services.html

STUDY

Our study aimed on running an exploratory analysis on mobile information supply and to investigate how users interact with information. We conducted a study following a passive observational design [23]. To understand how mobile devices supply information to their users we employed our measure in the wild. Only by conducting a large-scale study in the wild we have been able to observe participants in their natural ways of interacting with their usual content and pieces of information.

For our study we have been able to leverage an application which is publicly available on the Google Play Store, called AppDetox[2], for recruiting participants for our study. AppDetox is an app for users to study and control their own application usage. To ask for user consent we have used a two-button approach [19]. At time of the study AppDetox had about 140,000 total downloads and 24,000 active installations. We have been able to collect data from 1,385 users who agreed to submit their data in September 2017. People had been able to withdraw from the study at anytime through the app.

For employing our information measure we extended AppDetox to log two types of events: application switches and window content changes. All events have been timestamped UTC (we kept track of users' timezone offset to UTC). Whenever the window content has changed, we used our measure described above to keep track of the information supply, i.e. amount of new information the content provided. In addition we also kept track of the new content's number of raw bytes (i.e. the uncompressed length of the text on screen). All data was stored locally on participants' smartphones. The app frequently uploaded the data securely to our servers.

If users type in text via the keyboard this will also change the text on screen, since most apps print what the user is typing (e.g. a messaging app showing the chat flow). This information will also considered as information supply, even though this information is not provided to the user but rather supplied by the user herself. However, we argue that users might reread their own typing, hence they will also consume this amount of information.

We conducted a pre-study with 14 participants over 7 days for proofing our concept and testing our implementation. Later we found that most of the characteristics we already found by analyzing this small sample yield true for the large sample, but effects showed up more clearly.

RESULTS

We discarded data of users with bogus data (e.g. implausible timestamps, maybe due to users changing their wall clock settings). Further, we removed data from users who had less

[2] https://play.google.com/store/apps/details?id=de.dfki.appdetox

than 7 days of data in total. Afterwards, we finally had a valid dataset of 792 participants in total. In addition to data cleaning we also reconstructed device usage sessions from the events we have logged. Instead of logging screen on and off events, we considered subsequent events to belong to the same session if the gap between those is less than 45 seconds, as proposed by van Berkel et al. [28].

Demographics

We did not ask our participants for demographic information within the app. However, as the AppDetox app is using Google Analytics we can provide some insights on our total population: There are 58% male and 42% female users of the application. 36% are aged 18-24, 44% 25-34, 14% 35-44, 4% 45-54, 1% 55-64, and 1% 5+. The majority of users resides in the USA (21%), followed by India (17%), Brasilia (5%) and Germany (4%). From our logging data we know that 70% of our users' devices use English language.

General Usage and Information Supply

Our participants in sum used 6,817 different apps. On average a single participant had about 449 sessions per day (SD: 222), while a single session on average lasts about 100 seconds (SD: 42). Within a single session the smartphone supplies on average about 1143 bytes of new information (SD: 987). This leads to an information supply of 13.55 bytes per seconds (SD: 14.75).

We found an expected positive correlation between the length of a session and its information supply ($r = .69, p < .05$). That is, the longer a session lasts the more information is supplied. However, this correlation's cause is unclear: we cannot tell whether people have more time and consume more pulled information or there is more pushed information and people need time to consume it.

Daily Information Supply

Since mobile app usage has been reported to differ over the course of a day (e.g. [2, 29]), we investigated the daily distribution of information supply. Figure 1 shows the average information a single screen supplies (y-axis), categorized by hour of day (x-axis). The amount of information supply is macro-averaged per user (to give equal weight to each user) and per hour of day. The box plots show the information supply for all users during a specific hour (graph omits outliers). The graph shows that in the morning hours (about 2am to 4am) the information supply is lower than during the rest of the day.

We found that information supply correlates to the general diurnal app usage [2], which could have been expected. However, while we know that apps are being used throughout the day, our finding suggests that apps also provide new pieces of information throughout the day.

Redundancy and Revisitation of Content

We investigated the ratio between the raw text being presented on the screen and its information supply. This allows to assess the degree to which users see new content on the screen compared to revisiting old content (e.g. revisiting tweets on Twitter which have already been seen).

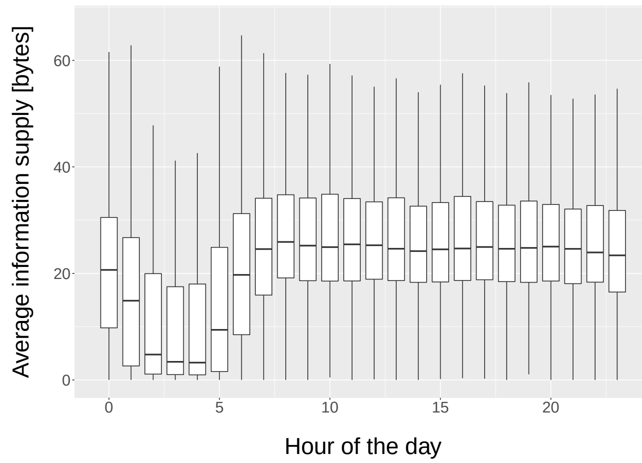

Figure 1: Average information supply over the course of a day.

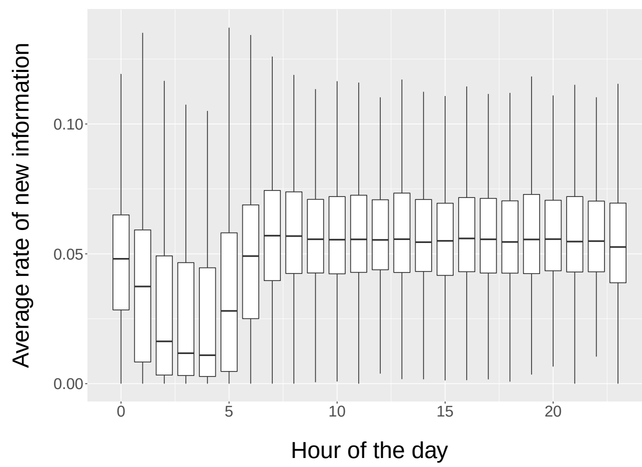

Figure 2: Average ratio between new and old information over the course of a day.

Figure 2 shows the ratio between new and old information a single screen provides on average (y-axis) categorized by hour of the day (x-axis). The ratio technically ranges from 0 (completely redundant information) to 1 (completely new information). However, approximating a value of 1 is almost impossible as described in the upcoming section, thus the y-axis only display the lower end of the ratio.

The graph shows that the share of newly supplied information is lower in the morning hours (2am to 4am), due to the fact that the redundancy is higher compared to the rest of the day. This suggests, that during this time the revisitation of content is higher. While it is obvious that there is much less newly generated content at night (since other people as sources of information are sleeping), it is surprising to see that participants nevertheless revisit old content.

This reveals that in the morning hours smartphones supply less new information (Figure 1) due to the fact that the revisitation is higher at that times (Figure 2).

Differences between Users

We also investigated the ratio of raw text and its information supply with regard to differences between users to study individual revisitation behavior. Thus, for each user we determined the ratio of new content a single screen provides on average.

Figure 3 shows the resulting frequency distribution with the different ratios of new information (x-axis), accumulated by users (y-axis). The graph shows a bell-shaped curve with a peak around a ratio of 0.05. Hence, many users share a similar information revisitation behavior with a rather low rate of new information (0.05). Only a very few participants are more efficient than the majority of users (ratio above 0.1).

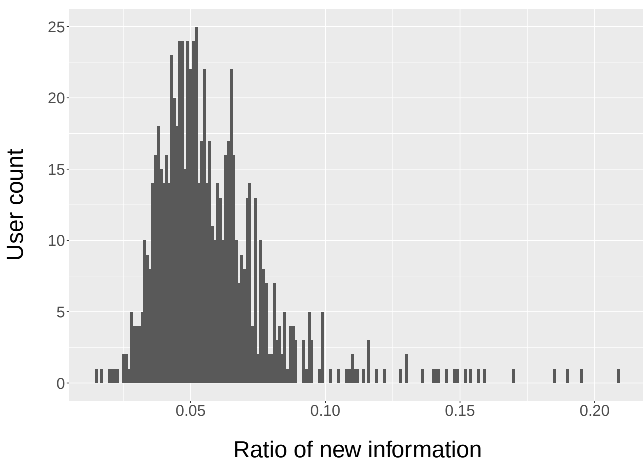

Figure 3: Accumulated amount of users with different ratios of new content.

A ratio of 1 would mean no redundancy in supplied text and, hence, no revisitation. This cannot be achieved in practical device usage, as apps will supply redundant content for example because of the ui structure and hardcoded labels. Furthermore, the values are low in general since language itself can already be compressed to a high degree (e.g. English has a redundancy of 75% [25]). However, the higher the ratio the more efficient a user's smartphone usage can be assessed thinking in terms of same text being presented less often. Our study suggests that only a few users have a more efficient information usage than the majority.

Two Types of Sessions

When considering the ratio count grouped by single sessions, as shown in Figure 4, we see two peaks (note that x-axis is log scale). The first peak is even before a rate of 0.01 which shows that a lot of sessions represent a rather inefficient smartphone usage, which could be based on usage sessions driven by a checking habit. However, the other peak shows that there are more efficient usage sessions too, as various sessions move around a ratio of 0.1.

This suggests that users have two types of sessions: one where they have a tendency to revisit content they have previously already seen, and another type where there is much more fresh content. This supports the checking habit [16], which has been described to also happen if no new content is available, hence, old content is revisited.

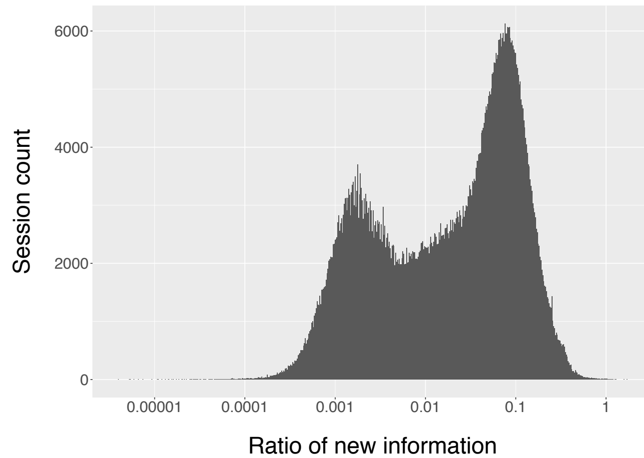

Figure 4: Accumulated amount of sessions with different ratios of new content (x-axis is log scale).

Information Supply within Sessions

We have also been interested in how information supply is distributed over the course of smartphone usage sessions. Therefore, we analyzed sessions over time from their begin ($t = 0$) to their end ($t = 1$). We use relative time frames for comparing sessions. Hence, the timestamps of events related to information supply have been interpolated within such session intervals.

Figure 5 shows a regression analysis of information supply within sessions. Each screen update is arranged to its relative point in time of the course of a session (x-axis) and assigned the average information supply (y-axis) to that time. The resulting regression curve shows that there is a higher information supply at the beginning of sessions, it declines quickly and remains low in the middle, and rises again at the end.

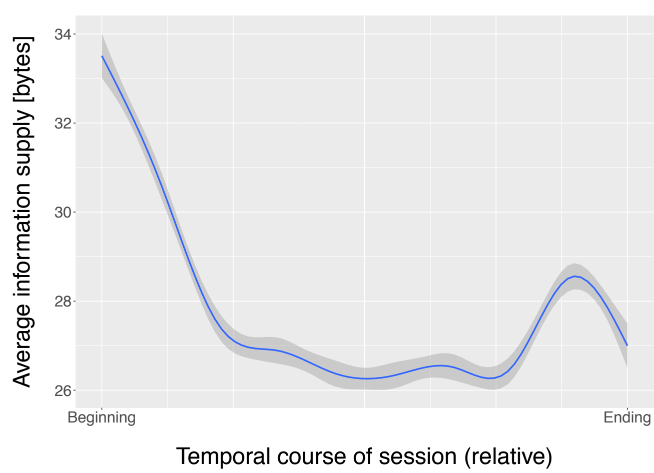

Figure 5: Information supply over the course of smartphone sessions (regression line with 95% confidence bounds).

The course of information supply within sessions has characteristics which can be described relying on information foraging theory [20]: While at the beginning of sessions the advent of new information stimulates users to use their devices (e.g. new notifications or an information need), one would expect this information to diminish after it has been consumed. However, for the ending of sessions one would expect that information supply further levels off. However, instead we see an increase at the end of sessions.

DISCUSSION

Our measure of information supply and the results of our large-scale study contribute to a better understanding of mobile interaction with information at hand, and provide some points for discussion.

13 Bytes per Second

Our study reveals that the bandwidth of smartphone displays for textual information is about 13 bytes per second.

It was found that during lexical decision tasks the human brain can process about 60 bits per second [14], which is 7.5 bytes per second. Hence, an average rate of 13 bytes per second might be an indication for information overload. However, as stated earlier, our measure cannot simply be transferred back to written text. Furthermore the standard deviation is rather high (14.75). Hence, at some points in time there might be a higher textual information supply on the screen than a user will be able to consume and at some points in time the information supply is rather low.

As like most studies conducted in-the wild, our study is limited in terms of explaining the causes of the effects we see in our information supply data. On the one hand a high information supply is an indication for an information overflow, while in truth it might represent an efficient usage session (e.g. reading a book). On the other hand a low information supply might be an indication for a rather inefficient usage session, while in truth it might represent the consumption of non-textual content we cannot quantify (e.g. watching a learning video), which results in a lower information supply per second.

However, this finding can still be used as a figure which can be taken into account for designing apps. It can be used in two ways: On the one hand, when supplying text to users, app designers should take care to not overload the user beyond overall 13 compressed bytes/sec, including the app itself but also the surrounding operating system. If app designers find a higher supply rate than we found on average, it is likely that the information of the text presented has not been consumed in total. On the other hand, app designers can evaluate the efficiency of different designs (e.g suppling different bytes/sec) while for example analyzing the resulting task completion times of single users.

Towards a Mobile Information Foraging Theory

We found that mobile information utilization can only partly be described with Pirolli and Card's information foraging theory [20]. If we think of a mobile device usage session we would have expected to have a high information supply at the beginning, which then levels until it reaches a minimum at the end. However, in our data (see Figure 5) we can find a peak at the sessions' beginning which then levels off, but in the end information supply rises again. The latter cannot be explained by the theory.

Pirolli and Card's [20] original information foraging theory does not take into account pushed information like notifications. That is, information foraging theory only describes the directed search and navigation for new information on the users intent. However, on smartphones there are many asynchronous sources for information, e.g. notifications available on the device and but also contextual cues which result in users taking actions on their smartphones. This behavior does not follow the idea of users discovering pieces of information, but users being asynchronously confronted with those. Thinking in terms of the original foraging theory, where mankind is searching for prey in a patch until the patch runs dry of available food, notifications could be more set into analogy as prey falling from the sky, like Cockaigne—a land of plenty.

It is surprising to see that at the end of sessions information supply increases (see Figure 5). One explanation might be that we also record those pieces of information users enter into smartphones themselves: it might happen that they type in a message, send it, and then end their session. Another explanation might be that after reading all the notifications available at the beginning of a session (while information suppy falls down), users start to take actions on their own intent to retrieve some information before they end their sessions.

While both explanations are reasonable, this phenomenon could also stimulate behavior related to smartphone addiction: after available information has been consumed in a session users might start taking actions which result in new content being retrieved. Then, before diving deeper into that content, users might decide to end their sessions before spending too much time on their phones. However, since we cannot know who of our 792 users are addicted to their smartphones, this might be an explanation for the reported phenomenon, but surely does not hold true for all of our participants.

CONCLUSION

This paper contributes a measure for studying mobile information supply. We presented a measure for quantifying the amount on textual information provided through smartphones by analyzing the content presented to users and applying concepts of information theory.

We found that the bandwidth of the communication channel between smartphones and their users is about 13 byte per second and 1143 bytes per device session. Within single sessions the information supply typically peaks at the beginning, which can be explained by incoming notifications and the user retrieving content. Then the information supply quickly levels off, which can be described by information foraging theory. However, surprisingly at the end of sessions we found another rise in information supply which can be described by the nature of smartphones. We also found that information supply is rather constant throughout the day while being at minimum during the early morning. From our data we have also been able to distinguish two types of sessions: on the one hand, we have sessions with a rather low share of new content, and on the other hand we have sessions with a higher share of new content. The first type of sessions reflects smartphone checking habit in cases where no new content is available.

Our community began to understand mobile device usage and its implications based on different levels like apps [2, 29, 11], web usage [26, 5, 6], and notifications [18, 21, 15]. This paper gives rise to a more fundamental understanding of mobile device usage related to information foraging theory [20]. We contribute our source code for others to run their own studies.[3] Our results include fairly high standard derivations which might result from different user groups, use cases and utilized hardware. Thus, our future work will be about investigating mobile information supply with regard to context, device characteristics, app type, differences between users as well as analyzing and evaluating image and video based screen content.

REFERENCES

1. Nikola Banovic, Christina Brant, Jennifer Mankoff, and Anind Dey. 2014. ProactiveTasks: The Short of Mobile Device Use Sessions. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 243–252. DOI: http://dx.doi.org/10.1145/2628363.2628380
2. Matthias Böhmer, Brent Hecht, Johannes Schöning, Antonio Krüger, and Gernot Bauer. 2011. Falling Asleep with Angry Birds, Facebook and Kindle: A Large Scale Study on Mobile Application Usage. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. ACM, New York, NY, USA, 47–56. DOI:http://dx.doi.org/10.1145/2037373.2037383
3. Barry Brown, Moira McGregor, and Eric Laurier. 2013. iPhone in Vivo: Video Analysis of Mobile Device Use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1031–1040. DOI: http://dx.doi.org/10.1145/2470654.2466132
4. Juan Pablo Carrascal and Karen Church. 2015. An In-Situ Study of Mobile App & Mobile Search Interactions. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2739–2748. DOI: http://dx.doi.org/10.1145/2702123.2702486
5. Karen Church, Mauro Cherubini, and Nuria Oliver. 2014. A Large-scale Study of Daily Information Needs Captured in Situ. *ACM Trans. Comput.-Hum. Interact.* 21, 2, Article 10 (Feb. 2014), 46 pages. DOI: http://dx.doi.org/10.1145/2552193
6. Karen Church and Nuria Oliver. 2011. Understanding Mobile Web and Mobile Search Use in Today's Dynamic Mobile Landscape. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. ACM, New York, NY, USA, 67–76. DOI:http://dx.doi.org/10.1145/2037373.2037385
7. L. Peter Deutsch. 1996. *DEFLATE Compressed Data Format Specification version 1.3*. RFC 1951. The Internet Engineering Task Force. 1–16 pages. https://www.ietf.org/rfc/rfc1951.txt
8. Don C Donderi. 2006. An Information Theory Analysis of Visual Complexity and Dissimilarity. *Perception* 35, 6 (2006), 823–835. DOI:http://dx.doi.org/10.1068/p5249
9. Denzil Ferreira, Jorge Goncalves, Vassilis Kostakos, Louise Barkhuus, and Anind K. Dey. 2014. Contextual Experience Sampling of Mobile Application Micro-usage. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 91–100. DOI: http://dx.doi.org/10.1145/2628363.2628367
10. Wai-Tat Fu and Peter Pirolli. 2007. SNIF-ACT: A Cognitive Model of User Navigation on the World Wide Web. *Hum.-Comput. Interact.* 22, 4 (Nov. 2007), 355–412. http://dl.acm.org/citation.cfm?id=1466607.1466608
11. Simon L. Jones, Denzil Ferreira, Simo Hosio, Jorge Goncalves, and Vassilis Kostakos. 2015. Revisitation Analysis of Smartphone App Use. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 1197–1208. DOI: http://dx.doi.org/10.1145/2750858.2807542
12. Juuso Karikoski and Tapio Soikkeli. 2013. Contextual Usage Patterns in Smartphone Communication Services. *Personal Ubiquitous Comput.* 17, 3 (March 2013), 491–502. DOI: http://dx.doi.org/10.1007/s00779-011-0503-0
13. Carolynne Lord, Mike Hazas, Adrian K. Clear, Oliver Bates, Rosalind Whittam, Janine Morley, and Adrian Friday. 2015. Demand in My Pocket: Mobile Devices and the Data Connectivity Marshalled in Support of Everyday Practice. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2729–2738. DOI: http://dx.doi.org/10.1145/2702123.2702162
14. Fermín Moscoso del Prado Martín. 2009. The thermodynamics of human reaction times. *arXiv preprint arXiv:0908.3170* (2009).
15. Abhinav Mehrotra, Veljko Pejovic, Jo Vermeulen, Robert Hendley, and Mirco Musolesi. 2016. My Phone and Me: Understanding People's Receptivity to Mobile Notifications. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1021–1032. DOI: http://dx.doi.org/10.1145/2858036.2858566

[3] https://github.com/C0d3GGz/MIG

16. Antti Oulasvirta, Tye Rattenbury, Lingyi Ma, and Eeva Raita. 2012. Habits Make Smartphone Use More Pervasive. *Personal Ubiquitous Comput.* 16, 1 (Jan. 2012), 105–114. DOI: http://dx.doi.org/10.1007/s00779-011-0412-2

17. Antti Oulasvirta, Sakari Tamminen, Virpi Roto, and Jaana Kuorelahti. 2005. Interaction in 4-second Bursts: The Fragmented Nature of Attentional Resources in Mobile HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '05)*. ACM, New York, NY, USA, 919–928. DOI: http://dx.doi.org/10.1145/1054972.1055101

18. Martin Pielot, Karen Church, and Rodrigo de Oliveira. 2014. An In-situ Study of Mobile Phone Notifications. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 233–242. DOI: http://dx.doi.org/10.1145/2628363.2628364

19. Martin Pielot, Niels Henze, and Susanne Boll. 2011. Experiments in App Stores - How to Ask Users for their Consent?. In *CHI '11 Workshop on Ethics, Logs and Videotape: Ethics in Large Scale Trials & User Generated Content*.

20. Peter Pirolli and Stuart Card. 1995. Information Foraging in Information Access Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '95)*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 51–58. DOI: http://dx.doi.org/10.1145/223904.223911

21. Alireza Sahami Shirazi, Niels Henze, Tilman Dingler, Martin Pielot, Dominik Weber, and Albrecht Schmidt. 2014. Large-scale Assessment of Mobile Notifications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3055–3064. DOI: http://dx.doi.org/10.1145/2556288.2557189

22. Alireza Sahami Shirazi, Niels Henze, Albrecht Schmidt, Robin Goldberg, Benjamin Schmidt, and Hansjörg Schmauder. 2013. Insights into Layout Patterns of Mobile User Interfaces by an Automatic Analysis of Android Apps. In *Proceedings of the 5th ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS '13)*. ACM, New York, NY, USA, 275–284. DOI: http://dx.doi.org/10.1145/2494603.2480308

23. William R Shadish, Thomas D Cook, and Donald T Campbell. 2002. *Experimental and quasi-experimental designs for generalized causal inference.* Houghton, Mifflin and Company.

24. C. E. Shannon. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27, 3 (July 1948), 379–423. DOI: http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x

25. Claude E Shannon. 1951. Prediction and entropy of printed English. *Bell Labs Technical Journal* 30, 1 (1951), 50–64.

26. Chad Tossell, Philip Kortum, Ahmad Rahmati, Clayton Shepard, and Lin Zhong. 2012. Characterizing Web Use on Smartphones. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 2769–2778. DOI: http://dx.doi.org/10.1145/2207676.2208676

27. Alexandre N. Tuch, Javier A. Bargas-Avila, Klaus Opwis, and Frank H. Wilhelm. 2009. Visual complexity of websites: Effects on users' experience, physiology, performance, and memory. *International Journal of Human-Computer Studies* 67, 9 (2009), 703 – 715. http://dx.doi.org/10.1016/j.ijhcs.2009.04.002

28. Niels van Berkel, Chu Luo, Theodoros Anagnostopoulos, Denzil Ferreira, Jorge Goncalves, Simo Hosio, and Vassilis Kostakos. 2016. A Systematic Assessment of Smartphone Usage Gaps. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4711–4721. DOI: http://dx.doi.org/10.1145/2858036.2858348

29. Kelly Widdicks, Oliver Bates, Mike Hazas, Adrian Friday, and Alastair R. Beresford. 2017. Demand Around the Clock: Time Use and Data Demand of Mobile Devices in Everyday Life. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 5361–5372. DOI: http://dx.doi.org/10.1145/3025453.3025730